4 Spectral theorem for self adjoint operators

Throughout the following section $\mathcal{H}$ is always a Hilbert space over $\mathbb{K} = \mathbb{C}$.

4.1 Motivation and preliminaries

diagonalization of self-adjoint matrix
$A = UDU^*$
D ... diagonal
U ... unitary

Let

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \in \mathbb{C}^{n \times n}$$

be a self adjoint matrix with eigenvalues $\lambda_1 < \lambda_2 \cdots < \lambda_n$. The orthogonal projections onto the corresponding eigenspaces are given by

$$E(\{\lambda_2\}) = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 0 & \\ & & & \ddots \\ & & & & 0 \end{pmatrix}$$

$$E(\{\lambda_1\}) := \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}, \quad \ldots, \quad E(\{\lambda_n\}) := \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & 1 \end{pmatrix}.$$

With these projections one can write

$$E = \sum_{k=1}^{n} \delta_{\lambda_k}$$

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 & & \\ & 0 & \\ & & \ddots \\ & & & 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 0 & \\ & & & \ddots \\ & & & & 0 \end{pmatrix} + \ldots + \lambda_n \begin{pmatrix} 0 & & \\ & \ddots & \\ & & 0 \\ & & & 1 \end{pmatrix}$$

$$A = \sum_{k=1}^{n} \lambda_k E(\{\lambda_k\}) = \boxed{\int_{\mathbb{R}} \mu \, dE(\mu) = A}$$

where the integral is with respect to the measure E which has point masses at $\lambda_1, \ldots, \lambda_n$. With the help of this measure one gets for any open interval $\Delta \subset \mathbb{R}$

$$E(\Delta) = \sum_{\lambda_k \in \Delta} E(\{\lambda_k\}) = \mathbb{1}_{\Delta}(A).$$

Goal: We want to show that for any $A = A^* \in \mathcal{L}(\mathcal{H})$ there exists a *spectral measure* E (which will be an orthogonal projection for each Borel set) such that

$$\boxed{A = \int_{\mathbb{R}} \mu \, dE(\mu)} = \int_{\sigma(A)} \mu \, dE(\mu).$$

Idea: Set $E(\Delta) = \mathbb{1}_{\Delta}(A)$ for any interval $\Delta \subset \mathbb{R}$. But how can $\mathbb{1}_{\Delta}(A)$ be understood and introduced? A function of an operator can be defined, if the function is a polynomial:

Definition 4.1. Let $A = A^* \in \mathcal{L}(\mathcal{H})$ and $p(t) = \sum_{k=0}^{n} a_k t^k$ be a polynomial on $\mathbb{R}$ with complex coefficients $a_0, \ldots, a_n$. Then $p(A)$ is defined by

$$\boxed{p(A) = \sum_{k=0}^{n} a_k A^k.}$$

$$A \in \mathcal{L}(\mathcal{H}) \Rightarrow A^k = \underbrace{A \cdot \ldots \cdot A}_{k \text{ times}} \in \mathcal{L}(\mathcal{H})$$

Lemma 4.2. Let $A = A^* \in \mathcal{L}(\mathcal{H})$ and let $p : \mathbb{R} \to \mathbb{C}$ be a polynomial. Then one has

$$\sigma(p(A)) = p(\sigma(A)) = \{p(\lambda) : \lambda \in \sigma(A)\}.$$

Proof:

"$\supset$" Let $\lambda \in \sigma(A)$ be fixed.

The polynomial $t \mapsto p(t) - p(\lambda)$ has a zero at $t = \lambda$.

$$\Rightarrow p(t) - p(\lambda) = (t-\lambda) \cdot \underbrace{q(t)}_{\text{another polynomial}} = q(t) \cdot (t-\lambda)$$

$$\Rightarrow p(A) - p(\lambda) = (A-\lambda) \cdot q(A) = q(A)(A-\lambda)$$

Assume that $p(\lambda) \in \rho(p(A))$

$$\Rightarrow I = (p(A) - p(\lambda)) \underbrace{(p(A) - p(\lambda))^{-1}}_{\in \mathcal{L}(\mathcal{H})} = (A-\lambda) \, q(A) \cdot (p(A) - p(\lambda))^{-1}$$

$$I = (p(A) - p(\lambda))^{-1} (p(A) - p(\lambda)) = (p(A) - p(\lambda))^{-1} q(A) \cdot (A - \lambda)$$

$\Rightarrow A - \lambda$ is bijective $\Rightarrow \lambda \in \rho(A)$ ↯ $\Rightarrow p(\lambda) \in \sigma(p(A))$ ✓

"$\subset$" Let $\mu \in \sigma(p(A))$. There exist $\lambda_1, \dots, \lambda_n \in \mathbb{C}$, $a \in \mathbb{C}$ s.t. $p(t) - \mu = a(t-\lambda_1)(t-\lambda_2) \dots (t-\lambda_n)$ $(*)$

$$\Rightarrow p(A) - \mu = a(A-\lambda_1) \dots (A-\lambda_n)$$

$\Rightarrow \exists k \in \{1, \dots, n\} : \lambda_k \in \sigma(A)$, as otherwise the operator in the line above would be invertible contradicting $\mu \in \sigma(p(A))$

$\overset{(*)}{\Rightarrow}$ $p(t) - \mu \,\big|_{t=\lambda_k} = 0 \Rightarrow \mu = p(\lambda_k) \in p(\sigma(A))$ $\square$

Recall: $\sigma(A)$ is compact for $A \in \mathcal{L}(\mathcal{H})$

4.2 The continuous functional calculus for self adjoint operators

Throughout this section we assume that A is a bounded and self-adjoint operator, i.e. $A = A^* \in \mathcal{L}(\mathcal{H})$. The goal is to define the operator $f(A)$ for any continuous function f. Denote by $C(\sigma(A))$ the set of all continuous functions $f : \sigma(A) \to \mathbb{C}$ equipped with the norm

$$\|f\|_\infty := \sup_{x \in \sigma(A)} |f(x)|, \qquad f \in C(\sigma(A)).$$

By $P(\sigma(A))$ we denote the space of all polynomials defined on $\sigma(A)$. By the Weierstrass approximation theorem (see e.g. [Werner, Satz VIII.4.7]) we have that $P(\sigma(A))$ is dense in $C(\sigma(A))$.

Theorem 4.3. Let $A = A^* \in \mathcal{L}(\mathcal{H})$. Then the map

$$P(\sigma(A)) \ni p \mapsto p(A) \in \mathcal{L}(\mathcal{H})$$

is linear and isometric, i.e. $\|p(A)\| = \|p\|_\infty$ for all $p \in P(\sigma(A))$, and hence it has a unique isometric (and thus bounded) linear extension $\Phi : C(\sigma(A)) \to \mathcal{L}(\mathcal{H})$, which has the following properties:

(a) Φ is multiplicative, i.e. $\Phi(fg) = \Phi(f)\Phi(g)$ for all $f, g \in C(\sigma(A))$.

(b) Φ is an involution, i.e. $\Phi(\overline{f}) = \Phi(f)^*$ for all $f \in C(\sigma(A))$.

Of course, the map Φ depends on the initially given operator A. We write

$$f(A) := \Phi(f), \qquad f \in C(\sigma(A)).$$

Due to the previous theorem we have $\|f(A)\| = \|f\|_\infty$ for all $f \in C(\sigma(A))$. The map Φ is called *continuous functional calculus* for A (the word continuous is associated to the fact that it applies to continuous functions).

Proof. For $p \in P(\sigma(A))$ define $\Phi_0(p) := p(A)$. Then

$$\|\Phi_0(p)\|^2 = \|p(A)\|^2 \underset{\text{see proof of Prop. 2.8}}{=} \|(p(A))^* p(A)\| \underset{p(A)^* = \bar{p}(A)}{=} \|\bar{p}(A)\, p(A)\|$$

$$= \|\underbrace{(\bar{p}p)(A)}_{=(p\cdot\bar{p})(A) = \dots = p(A)\cdot p(A)^*}\| \Rightarrow p(A) \text{ is normal}$$

$$\overset{\text{Prop. 2.8}}{=} \inf_{n \in \mathbb{N}} \|((\bar{p}\cdot p)(A))^n\|^{1/n} \overset{\text{Theorem 2.7}}{=} \sup\{|\lambda| : \lambda \in \sigma((\bar{p}\cdot p)(A))\}$$

$$\overset{\text{Lemma 4.2}}{=} \sup\{|p(\mu)|^2 : \mu \in \sigma(A)\} = \|p\|_\infty^2$$

$\Rightarrow \quad \|\Phi_0(p)\| = \|p\|_\infty \qquad \forall p \in P(\sigma(A))$

$\Rightarrow$ Φ_0 is an isometry

Since $P(\sigma(A))$ are dense $C(\sigma(A))$ by the Weierstrass approximation theorem, Φ_0 can be extended by continuity to an isometry $\Phi : C(\sigma(A)) \to \mathcal{L}(\mathcal{H})$ by

$\Phi(f) := \lim_{n\to\infty} \Phi_0(p_n)$, where the limit in $\mathcal{L}(\mathcal{H})$ and $(p_n) \subset P(\sigma(A))$ with $\|p_n - f\|_\infty \to 0$

In order to prove (b) [the proof of (a) is similar], take for $f \in C(\sigma(A))$ a sequence $(p_n) \subset P(\sigma(A))$ s.t. $\|p_n - f\|_\infty \xrightarrow[n\to\infty]{} 0$.

Then we have $\quad \overline{p_n}(A) = \underline{\Phi_0(\overline{p_n})} = (p_n(A))^* = \underline{(\Phi_0(p_n))^*}$

$$\Rightarrow \underline{\Phi(f)^*} = \left(\lim_{n\to\infty} \Phi(p_n)\right)^* = \lim_{n\to\infty} \Phi(p_n)^* = \lim_{n\to\infty} \Phi_0(\overline{p_n}) = \underline{\Phi(\overline{f})} \checkmark$$

$\square$

Proposition 4.4. Let $A = A^* \in \mathcal{L}(\mathcal{H})$. Then the following holds for all $f, g \in C(\sigma(A))$.

(i) $f(A)g(A) = g(A)f(A)$.

(ii) If $f(t) \geq 0$ for all $t \in \sigma(A)$, then $f(A) \geq 0$ in the sense of self adjoint operators (i.e. $(f(A)x, x) \geq 0$ for all $x \in \mathcal{H}$).

(iii) $f(A)$ is a normal operator and $f(A) = f(A)^*$ if and only if f is real-valued.

(iv) $Ax = \lambda x$ implies $f(A)x = f(\lambda)x$.

Beweis. See exercises. □

Theorem 4.5 (Spectral mapping theorem)**.** Let $A = A^* \in \mathcal{L}(\mathcal{H})$. Then one has for all $f \in C(\sigma(A))$

$$\sigma(f(A)) = f(\sigma(A)) = \{f(\lambda) : \lambda \in \sigma(A)\}$$

Proof:

"$\subset$" Let $\mu \notin f(\sigma(A))$. $\Rightarrow$ $g(t) := \frac{1}{f(t)-\mu}$ fulfills $g \in C(\sigma(A))$ and $g \cdot (f-\mu) = 1_{\sigma(A)}$

$$\Rightarrow I = \Phi(1_{\sigma(A)}) = \Phi(g \cdot (f-\mu)) \overset{(a)}{=} \Phi(g) \cdot \Phi(f-\mu) = g(A) \cdot (f(A)-\mu) = (f(A)-\mu)\, g(A)$$

$\Rightarrow$ $f(A)-\mu$ is bijective, i.e. $\mu \in \rho(f(A))$ ✓

"$\supset$" Let $\mu = f(\lambda)$ for $\lambda \in \sigma(A)$.

Let $(p_n) \subset P(\sigma(A))$ s.t. $\|p_n - f\|_\infty < \frac{1}{n}$.

$\Rightarrow$ $|f(\lambda) - p_n(\lambda)| < \frac{1}{n}$ and $\|f(A) - p_n(A)\| = \|\Phi(f-p_n)\| = \|f-p_n\|_\infty < \frac{1}{n}$

By Lemma 4.2 we have $\sigma(p_n(A)) = p_n(\sigma(A))$. By Prop. 3.11 (which actually is true for all normal operators!) there exist for any $n \in \mathbb{N}$: $x_n \in \mathcal{H}$: $\|x_n\| = 1$ and $\|(p_n(A) - p_n(\lambda))x_n\| < \frac{1}{n}$

$$\|(f(A)-\underset{\substack{\| \\ f(\lambda)}}{\mu})x_n\| = \|(f(A)-p_n(A)+p_n(A)-p_n(\lambda)+p_n(\lambda)-f(\lambda))x_n\|$$

$$\leq \underbrace{\|(f(A)-p_n(A))x_n\|}_{\leq \underbrace{\|f(A)-p_n(A)\|}_{<\frac{1}{n}} \cdot \underbrace{\|x_n\|}_{=1}} + \underbrace{\|(p_n(A)-p_n(\lambda))x_n\|}_{<\frac{1}{n}} + \underbrace{\|(p_n(\lambda)-f(\lambda))x_n\|}_{\underbrace{|f(\lambda)-p_n(\lambda)|}_{<\frac{1}{n}} \cdot \underbrace{\|x_n\|}_{=1}}$$

$$< \frac{3}{n} \longrightarrow 0, \; n\to\infty$$

By the definition of $\rho(A)$ this is only possible, if $\mu = f(\lambda) \in \sigma(f(A))$ $\square$